JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1992, 58, 513-525 NUMBER 3 (NOVEMBER)

CHOICE WITH DELAYED AND PROBABILISTIC REINFORCERS: EFFECTS OF VARIABILITY, TIME BETWEEN TRIALS, AND CONDITIONED REINFORCERS

JAMES E. MAZUR AND ANGELA ROMANO

SOUTHERN CONNECTICUT STATE UNIVERSITY

In a discrete-trials procedure with pigeons, a response on a green key led to a 4-s delay (during which green houselights were lit) and then a reinforcer might or might not be delivered. A response on a red key led to a delay of adjustable duration (during which red houselights were lit) and then a certain reinforcer. The delay was adjusted so as to estimate an indifference point—a duration for which the two alternatives were equally preferred. Once the green key was chosen, a subject had to continue to respond on the green key until a reinforcer was delivered. Each response on the green key, plus the 4-s delay that followed every response, was called one "link" of the green-key schedule. Subjects showed much greater preference for the green key when the number of links before reinforcement was variable (averaging four) than when it was fixed (always exactly four). These findings are consistent with the view that probabilistic reinforcers are analogous to reinforcers delivered after variable delays. When successive links were separated by 4-s or 8-s "interlink intervals" with white houselights, preference for the probabilistic alternative decreased somewhat for 2 subjects but was unaffected for the other 2 subjects. When the interlink intervals had the same green houselights that were present during the 4-s delays, preference for the green key decreased substantially for all subjects. These results provided mixed support for the view that preference for a probabilistic reinforcer is inversely related to the duration of conditioned reinforcers that precede the delivery of food.

Key words: probability of reinforcement, delay of reinforcement, choice, conditioned reinforcement, key peck, pigeons

Rachlin, Logue, Gibbon, and Frankel (1986) proposed that probabilistic reinforcers (reinforcers delivered with a probability of less than 1.0) are functionally equivalent to delayed reinforcers. That is, when a reinforcer is delivered with a probability of less than 1.0, the effect on a subject's behavior is the same as if the reinforcer were delayed. To understand the reasoning of Rachlin and his colleagues, consider a discrete-trial situation in which a subject must choose between a certain reinforcer (i.e., one delivered with a probability of 1.0) and a probabilistic reinforcer (e.g., one delivered with a probability of .25). The probabilistic alternative will sometimes deliver a reinforcer the first time it is chosen, but more often it will deliver a reinforcer only after it has been chosen several times. With a probability of .25, a reinforcer will be delivered, on average, once for every four choices of the probabilistic alternative. Rachlin et al. suggested that because of the time needed to complete these additional trials, and because of any intertrial interval (ITI) that may occur between trials, the probabilistic reinforcer is functionally equivalent to a delayed reinforcer.

More specifically, Rachlin et al. (1986) suggested that the following equation describes the relation between probabilistic and delayed reinforcers:

$$D = \frac{c + t}{p} - t, \tag{1}$$

where D is the expected delay to a reinforcer delivered with a probability of p, t is the duration of the ITI, and c is time from the start of a trial until either a reinforcer is delivered or, on trials without reinforcement, until the trial ends and the ITI begins. Therefore c includes the time it takes a subject to make the choice response, plus any delay that may be imposed between the response and reinforcement. For example, suppose c is 2 s, t is 10 s, and p is .25. On average, it will take four trials

This research was supported by Grant MH 38357 from the National Institute of Mental Health. We thank Tanya Andrievsky, Jennifer Taft, and Tamara Wolfson for their help in conducting the experiment. Correspondence concerning this article should be addressed to James E. Mazur, Psychology Department, Southern Connecticut State University, New Haven, Connecticut 06515.

to obtain a reinforcer, so Equation 1 states that D, the expected time to reinforcement, is 38 s (four trials of 2 s each, separated by three ITIs of 10 s each). The theory of Rachlin et al. therefore makes a specific prediction for this example: This probabilistic reinforcer will be equally preferred to a reinforcer delivered with certainty after a delay of 38 s.

Rachlin et al. (1986) obtained some support for their theory of probabilistic reinforcement in an experiment with human subjects in a hypothetical gambling situation. Rachlin, Raineri, and Cross (1991) found further evidence for the correspondence between probabilistic and delayed reinforcers in another study with human subjects. However, in a series of experiments with pigeons, Mazur (1989) found less support for the predictions of Equation 1. In these experiments, pigeons repeatedly chose between a probabilistic reinforcer (one that might or might not be delivered after a 5-s delay) and a certain reinforcer that was always delivered after an adjusting delay. This delay was adjusted over trials to find an *indifference point*—a delay at which the two alternatives were chosen equally often. To put it simply, the indifference point was a direct measure of D, a delay for the certain reinforcer that made it equivalent (as judged by the pigeons' choice behavior) to the probabilistic reinforcer.

The estimates of D obtained in these experiments posed problems for Equation 1 in two respects. First, for a choice between a certain reinforcer and a probabilistic reinforcer, Equation 1 predicts that preference for the probabilistic reinforcer should decrease as the ITI increases. However, Mazur (1989) found no systematic effects on choice as the ITI was varied. Second, the obtained estimates of D were consistently shorter than predicted by Equation 1, even when t, the duration of the ITI, was removed from this equation (thus yielding $D = c/p$). Mazur proposed that the basic idea of an equivalence between probabilistic reinforcers and delayed reinforcers might be correct, but that probabilistic reinforcers might be analogous to reinforcers delivered after variable rather than fixed delays. After all, the time between a choice of the probabilistic alternative and the delivery of a reinforcer is indeed variable and unpredictable: A reinforcer might be delivered on the first trial or only after many trials. Because animals show a strong preference for reinforcers delivered after variable rather than fixed delays (Cicerone, 1976; Mazur, 1984; Rider, 1983), this could explain why the estimates of D were much shorter than predicted by Equation 1 (indicating a stronger preference for the probabilistic reinforcers than predicted by the equation).

If probabilistic reinforcers are equivalent to reinforcers delivered after variable delays, how can we predict D, the duration of an equivalent *fixed* delay? Based on its success in describing the results of previous studies on variable delays (e.g., Mazur, 1984; Mazur, Snyderman, & Coe, 1985), Mazur (1989) suggested using the following equation:

$$V = \sum_{i=1}^{n} P_i\left(\frac{1}{1 + KD_i}\right). \quad (2)$$

V is the value of a reinforcer delivered after a variable delay, where there are n different possible delays to reinforcement. *Value* refers to the reinforcer's ability to sustain choice responses. P_i is the probability that a delay of D_i seconds will occur on any given trial. (Note the distinction between P_i, the probability that a certain delay will occur, and p in Equation 1, the probability of reinforcement on a given trial.) K is a free parameter that determines how rapidly V declines with increasing values of D_i. Mazur (1984) obtained fairly good predictions if K was set equal to 1.

Mazur (1989) used Equation 2 to predict indifference points for his experiments on probabilistic reinforcement in the following way. In his procedure, each choice of the probabilistic alternative was made by one peck on a red key; the key was usually lit for about 1 s before a subject pecked it. This was followed by a 5-s delay in which a red houselight was lit, and then the reinforcer might or might not be delivered. Therefore, for each trial on which the red key was pecked, the subject spent about 6 s in the presence of red stimuli associated with the probabilistic alternative. Mazur defined D_i as the total time spent in the presence of these red stimuli prior to each reinforcer delivery. Thus, D_i equaled 6 s for all cases in which the reinforcer was delivered after one trial of the probabilistic alternative, 12 s for cases in which it was delivered after two trials, and so on. These values of D_i were used in Equation 2, weighted by the appropriate values of P_i, to obtain a value of V for the probabilistic alternative. To predict the delay for an equally preferred certain reinforcer, this value of V was then used in Equation 2 with $n = 1$ and $P_i = 1$ (because the certain reinforcer was delivered each time it was chosen), and the equation was then solved to yield D.

The predictions that Mazur (1989) obtained from Equation 2 were much better than those from Equation 1, thus lending credence to the idea that probabilistic reinforcers are analogous to reinforcers delivered after variable delays. Mazur (1991) found additional support for Equation 2 in a study in which delay and probability of reinforcement were systematically varied. However, none of these studies specifically tested the assumption that what made Equation 2 more accurate than Equation 1 was the variability inherent in the probabilistic reinforcers. One purpose of the present experiment was to test this assumption more directly, by observing what happens to pigeons' choices when the variability of the probabilistic alternative is removed, thereby making it more similar to a reinforcer delivered after a fixed delay (as explained in more detail below). Subjects' choices were measured with the same type of adjusting-delay procedure used by Mazur (1989, 1991).

A second purpose of this experiment was to examine how the time between trials, and the stimuli present during these times, affected choice. As already mentioned, Mazur (1989) found that the duration of the ITI did not affect choice between a probabilistic reinforcer and a certain reinforcer. He speculated that this might be because the stimuli present during the ITI (white houselights) were not specifically associated with either alternative, and therefore did not contribute to D_i for either alternative. Mazur (1991) proposed that the stimuli associated with the probabilistic alternative (red keylight and houselight) were conditioned reinforcers because they were occasionally paired with food, and that preference for these reinforcers should decrease as their durations increase with no increase in primary reinforcement (cf. Dunn & Spetch, 1990; Spetch, Belke, Barnet, Dunn, & Pierce, 1990). He tested this idea by keeping the houselight red for 60 s (i.e., for the entire ITI), instead of only 5 s, on probabilistic trials that ended without reinforcement. If this increase in the duration of the red houselight served to increase D_i in Equation 2, this should produce a reduced preference for the probabilistic alternative. Preference decreased for 2 subjects, but showed no systematic change for the other 2.

To try to clarify these ambiguous results, the present experiment varied both the time between trials and the stimuli present during this time. However, the "trials" of this experiment were arranged in a different way than in the studies already mentioned. In the previous studies (both those with pigeons and those with human subjects), the time between a first choice of the probabilistic alternative and when it finally delivered a reinforcer was uncertain for two reasons—first, because of the probabilistic nature of the schedule itself, and second, because choices of the certain reinforcer were often intermixed with choices of the probabilistic alternative. That is, if one choice of the probabilistic alternative ended without reinforcement, a subject might choose the certain alternative rather than the probabilistic alternative on the next trial. In fact, in Mazur's (1989, 1991) experiments, some trials were forced trials in which the subject had to choose the certain alternative. In other words, many different possible sequences of events might follow a nonreinforced choice of the probabilistic alternative.

The present experiment used a procedure that was intended to reduce this complexity. If the first choice of the probabilistic alternative was not reinforced, the subject was then forced to choose the probabilistic alternative again and again until it finally delivered a reinforcer. The probabilistic alternative can therefore be considered a chained schedule in which the subject had to complete one or more links before a reinforcer was presented. Each *link* of the probabilistic alternative began with the illumination of a green keylight (the color always associated with the probabilistic alternative), followed by a response on the green key, a 4-s delay (during which the green houselights were lit), and then possible reinforcement. To avoid confusion, one *trial* of the probabilistic alternative will refer to a complete sequence of one or more such links that eventually ended with reinforcement. In some conditions, if the link did not end in reinforcement, the green key was immediately reilluminated to start the next link; in other conditions, an interlink interval (ILI) separated successive links of the probabilistic schedule. To test the predictions of Equation 2, in some conditions a white houselight was present during the ILIs, whereas in other conditions a green houselight (the stimulus associated with the probabilistic alternative) was present. If Mazur's (1991) analysis is correct, preference for the probabilistic alternative should decrease with the green houselights, because the increased duration of these houselights should increase D_i in Equation 2.

To examine the hypothesis that the variability of the probabilistic alternative affects preference for this alternative, the number of links required for reinforcement was constant in some conditions and variable in others. That is, in *fixed* conditions, the probabilistic alternative always delivered a reinforcer after *exactly* four links (presumably making the probabilistic alternative similar to a fixed delay before reinforcement). In *variable* conditions, the probabilistic reinforcer was delivered after an unpredictable number that only *averaged* four links (presumably making the probabilistic alternative similar to a variable delay before reinforcement). Because no reinforcer was ever delivered after only one, two, or three links in the fixed conditions, Equation 2 predicts that preference for the probabilistic alternative should be substantially lower than in the variable conditions. The predictions of Equation 2 were further tested by using two different distributions to schedule the probabilistic reinforcers in variable conditions.

METHOD

Subjects

Four White Carneau pigeons were maintained at about 80% of their free-feeding weights. All had previous experience with a variety of experimental procedures.

Apparatus

The experimental chamber was 30 cm long, 30 cm wide, and 33 cm high. Three response keys, each 1.8 cm in diameter, were mounted in the front wall of the chamber, 20.5 cm above the floor. A force of approximately 0.1 N was required to operate each key, and each effective response produced a feedback click. Each key could be transilluminated with lights of different colors. A hopper below the center key provided controlled access to mixed grain, and when grain was available, the hopper was illuminated with a 2-W white light. Eight 2-W lights (two white, two red, two green, and two blue) were mounted above the wire-mesh ceiling of the chamber. The chamber was enclosed in a sound-attenuating box containing a ventilation fan. All stimuli were controlled and responses recorded by an IBM-compatible personal computer using the Medstate® programming language.

Procedure

The experiment consisted of 14 conditions, each of which remained in effect for 14 or more sessions. Each session included several choice trials, on which the subject chose between the left key, which was transilluminated with green light, and the right key, which was transilluminated with red light. The green key will be called the *standard* key, because the schedule of reinforcement on this key remained constant throughout a condition. The red key will be called the *adjusting* key, because it included a delay to reinforcement that was systematically increased and decreased throughout each condition, as described below. This procedure was used to estimate an indifference point—a duration of the adjusting delay at which the two alternatives, standard and adjusting, were chosen about equally often. Each session also included several forced trials, on which only one side key, red or green, was illuminated, and pecks on the dark key had no effect.

To explain the details of the different conditions, the procedure for one type of condition—the fixed condition with no ILI—is described first. This description is followed by briefer explanations of the changes in this procedure that were implemented in the other conditions. Table 1 summarizes the schedules in effect in each of the 14 conditions.

Fixed conditions with no ILI (Conditions 1, 4, and 8). Each session lasted for 56 trials or for 60 min, whichever came first. Each block of four consecutive trials consisted of two forced trials followed by two choice trials. Each trial was preceded by an ITI, during which only the white houselights were lit. After the ITI, the center key was illuminated with white light to start a trial. A single peck on the center key was required to begin the choice period. The purpose of this center peck was to make it more likely that the subject's head was equidistant from the two side keys when the choice period

Table 1

Order of conditions and mean adjusting delays for each subject. All durations are in seconds. ILI = interlink interval, and ILI color is the houselight color during the ILI.

Condition	Distribution of links	ILI duration	ILI color	Mean adjusting delay Bird 1	Bird 2	Bird 3	Bird 4
1	Fixed	none	—	22.31	22.12	24.05	17.38
2	Exponential	none	—	9.50	8.72	13.50	8.81
3	Exponential	4 s	White	8.40	6.43	14.83	11.22
4	Fixed	none	—	27.03	26.67	31.19	16.74
5	Fixed	4 s	White	22.75	27.43	44.54	26.93
6	Fixed	4 s	Green	28.01	39.57	44.80	29.24
7	Fixed	4 s	White	22.31	35.26	43.59	27.36
8	Fixed	none	—	21.67	27.52	33.74	16.83
9	Rectangular	none	—	18.07	16.81	32.48	14.22
10	Exponential	none	—	8.07	10.10	11.19	7.81
11	Fixed	8 s	White	15.61	26.20	38.58	32.11
12	Fixed	8 s	Green	29.29	38.93	57.52	34.51
13	Fixed	8 s	White	30.71	29.73	43.51	34.31
14	One 16-s link	none	—	18.86	17.10	24.19	19.38

began. On choice trials, a peck on the center key darkened this key and illuminated both the left (green) key and the right (red) key.

Figure 1 illustrates the consequences of choosing either the green or the red key. One peck on the green key constituted a choice of the standard alternative, which in the fixed conditions always consisted of exactly four links, each beginning with a peck on the green key. The first link began with the choice response on the green key, which darkened both side keys and caused the houselight color to change from white to green for 4 s. At the end of this 4-s delay, the houselight color changed back to white, and the green key was again illuminated. A peck on the green key began the second link, which was the same as the first—a 4-s delay during which the green houselights were lit. The same sequence was followed for the remaining links. Thus each link consisted of a peck on the green key followed by a 4-s delay with green houselights. In the fourth link, the 4-s delay was followed by 3 s of access to food, during which all lights were off except the white lights illuminating the grain. After this reinforcement period, a 20-s ITI began, during which the white houselights were lit. Then the white center key was lit to start the next trial.

If the red key was pecked during the choice period, both keylights were extinguished and the adjusting delay began, during which the red houselights were lit. (The procedure for determining the duration of the adjusting delay is described below.) Following the adjusting delay, the red houselights were extinguished, and food was presented for 3 s, followed by an ITI during which the white houselights were on. The total time from a choice response until the end of the ITI will be called the *trial*

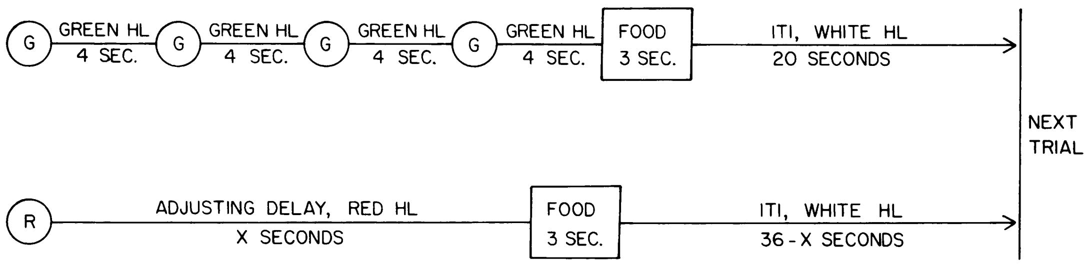

Fig. 1. For conditions with a fixed number of green-key links and no ILI, the consequences of choosing either the green key or the red key are shown. Each circle represents a lighted green (G) or red (R) response key, and one peck on the key was required to proceed with the sequence. (HL = houselight.)

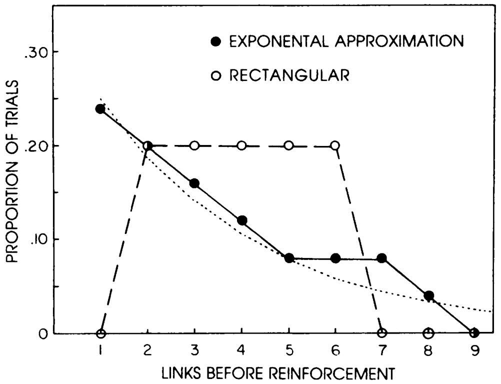

Fig. 2. The proportions of reinforcers delivered after different numbers of green-key links are shown for the exponential and rectangular distributions used in this experiment. The dotted line shows the proportions for an exact exponential distribution in which the conditional probability of reinforcement for each link was .25.

duration. On each adjusting trial, the duration of the ITI was set so that the trial duration was 39 s. This procedure was followed in order to keep the trial durations approximately the same on standard and adjusting trials. The trial duration on standard trials was slightly longer—39 s plus the time a subject needed to peck the green key during the second, third, and fourth links.

The procedure on forced trials was the same as on choice trials, except that only one side key was lit, red or green, and a peck on the lit key led to the appropriate schedule. A peck on the opposite (dark) key had no effect. Of every two forced trials, one involved the red key and the other the green key. The order of the red and green forced trials varied randomly.

At the start of every condition, the duration of the adjusting delay was set at 0 s. After every two choice trials, the delay for the adjusting key was increased by 1 s if the adjusting key was chosen on both choice trials, decreased by 1 s if the standard key was chosen on both choice trials, and remained unchanged if each was chosen once. In all three cases, this adjusting delay remained in effect for the next block of four trials. At the start of the second and all subsequent sessions of a condition, the adjusting delay was determined by the above rules as if it were a continuation of the preceding session.

Variable conditions with no ILI (Conditions 2, 9, and 10). The procedure in these conditions was the same as in the fixed conditions, except that the number of links for the standard alternative varied from trial to trial. In Conditions 2 and 10, the average number of links per trial was four, but the actual number of links varied from one to eight, based on a pseudorandom schedule that was designed to approximate an exponential distribution. In Figure 2, the filled circles show that the probability of having only one link before reinforcement was .24, the probability of having exactly two links was .20, and so on. The dotted line in Figure 2 shows the probabilities for a perfect exponential distribution in which the conditional probability of reinforcement for each successive link is .25. To limit the number of possible links to eight, the actual schedule departed slightly from this ideal function, but it was designed to approximate a situation in which each link of the green key had a roughly equal conditional probability of reinforcement; it will be called the *exponential* schedule.

The open circles in Figure 2 show the *rectangular* distribution of links that was used in Condition 9. In this condition, the number of required links varied from two to six, each occurring with a probability of .20. As with the fixed and exponential schedules, the average number of links before reinforcement with the rectangular schedule was four. This condition was included to determine whether subjects' choices would be sensitive not only to variability per se but to the shape of the variable distribution.

In all the variable conditions, the ITI after standard trials was 20 s, regardless of the number of links. This meant that trial durations varied with the number of links. However, as in the fixed conditions described above, the *average* trial duration on standard trials remained just slightly longer than the 39-s trial durations on adjusting trials, again due to the time needed to peck the green key at the start of each link.

Fixed conditions with white ILIs (Conditions 5, 7, 11, and 13). These conditions were similar to the fixed conditions without ILIs, except that after every link without reinforcement (i.e., the first three links), the 4-s delay with green houselights was followed by an ILI during which the white houselights were lit. The duration of each ILI was 4 s in Conditions 5 and 7 and 8 s in Conditions 11 and 13. At the end of each ILI, the green key was lit to begin the next link. The purpose of these conditions was to examine how and if preference for the standard key would change when there was more time between links, time during which the green houselights associated with the standard alternative were not present.

The presence of three 4-s ILIs added 12 s to the total duration of each standard trial, and the presence of three 8-s ILIs added 24 s to each standard trial. To compensate for these increased trial durations, the durations of adjusting trials were changed from 39 s to 51 s in conditions with 4-s ILIs and to 63 s in conditions with 8-s ILIs. In addition, because of the longer trial durations, sessions ended after a maximum of 48 trials in all conditions with 8-s ILIs.

Fixed conditions with green ILIs (Conditions 6 and 12). These conditions were identical in every way to the conditions with white ILIs, except that the green houselights were lit during the ILIs. ILI duration was 4 s in Condition 6 and 8 s in Condition 12. It may seem inappropriate to state that these conditions had ILIs at all, because there was no change in stimuli between the 4-s "delay" and the 4-s or 8-s "ILI" that followed. For instance, with the 8-s ILIs of Condition 12, the four pecks on the green key were followed by green houselights that lasted for 12 s, 12 s, 12 s, and 4 s, respectively. However, we have referred to these as conditions with green ILIs to emphasize their similarity to the conditions with white ILIs. The purpose of these conditions was, of course, to determine whether changing the color of the ILI stimulus would affect preference for the standard alternative.

Variable condition with 4-s white ILIs (Condition 3). This condition was similar to the fixed conditions with 4-s white ILIs, except that the number of links before reinforcement varied according to the exponential distribution represented by the filled circles in Figure 2.

Condition with a single, fixed delay (Condition 14). In this condition, a choice of the standard alternative led to a single delay of 16 s (with green houselights present), followed by a 3-s reinforcer. The purpose of this condition was to measure any possible bias for or against the adjusting key (see Mazur, 1984, 1986). If there were no key bias, the mean adjusting delay should approximately equal the standard delay of 16 s. ITIs were 20 s after standard trials, and as in all other conditions, ITIs after adjusting trials were chosen to equate trial durations for the two alternatives.

Criteria for terminating conditions. Conditions 6 through 10 lasted for a minimum of 14 sessions, and all other conditions lasted for a minimum of 20 sessions. After the minimum number of sessions, a condition was terminated for each subject individually when several stability criteria were met. To assess stability, each session was divided into two 28-trial blocks (two 24-trial blocks for conditions with only 48 trials), and for each block the mean delay on the adjusting key was calculated. The results from the first two sessions of a condition were not used, and the condition was terminated when the following three criteria were met, using the data from all subsequent sessions: (a) Neither the highest nor the lowest single-block mean of a phase could occur in the last six blocks of a condition. (b) The mean adjusting delay across the last six blocks could not be the highest or the lowest six-block mean of the condition. (c) The mean delay of the last six blocks could not differ from the mean of the preceding six blocks by more than 10% or by more than 1 s (whichever was larger).

RESULTS

The right side of Table 1 shows the mean adjusting delays for each subject in each condition. These mean adjusting delays will be referred to as the *indifference points*. These means, and all other analyses, were based on data from the six half-session blocks that satisfied the stability criteria.

Condition 14 tested subjects for any possible bias for the adjusting key. In this condition, choice of the standard key led to reinforcement after one 16-s delay. If there was no bias, the mean adjusting delay would also be approximately 16 s. However, the indifference points were greater than 16 s for all 4 subjects, and the group mean was 19.9 s, indicating some bias for the adjusting key. This bias should be taken into account when interpreting the results from other conditions.

Figure 3 shows the indifference points from the conditions with no ILIs. These results are from three conditions with a fixed distribution of links, two with an exponential distribution, and one with a rectangular distribution. For

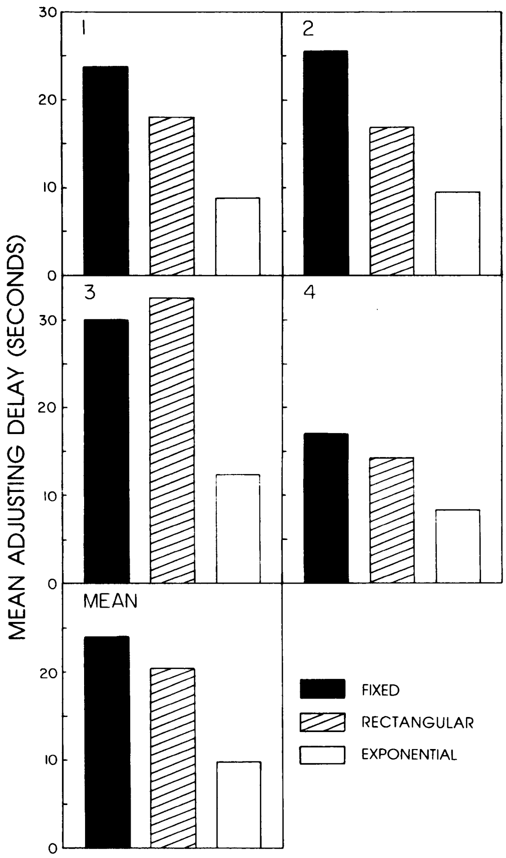

Fig. 3. The mean adjusting delays for each subject and for the group are shown for conditions with three different distributions of links and no ILIs.

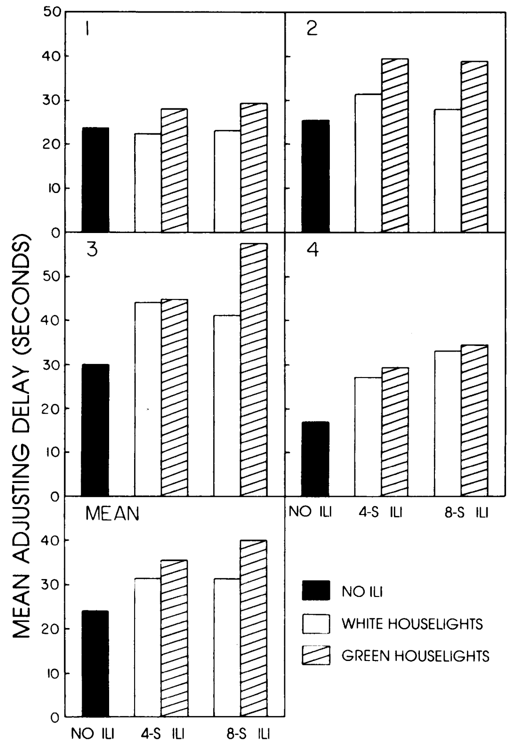

Fig. 4. The mean adjusting delays for each subject and for the group are shown for fixed conditions with and without ILIs.

the fixed and exponential distributions, the results are averaged across replications. Figure 3 shows that the indifference points were generally longest with the fixed distribution, shorter with the rectangular distribution, and shortest with the exponential distribution. The only exception was for Subject 3 in the rectangular condition. A one-way repeated-measures analysis of variance conducted on the results shown in Figure 3 found a statistically significant effect of the type of distribution, $F(2, 6) = 16.72$, $p < .01$. These results are consistent with Equation 2, which predicts that the fixed distribution of links should be least preferred (thus having the longest indifference points), and the exponential distribution should be most preferred. To derive more specific predictions from Equation 2, K was set equal to 1, and a response latency of 1.5 s was included for each peck on the green key at the start of a link (a value roughly equal to the mean latencies in these conditions, as presented below). With these values, Equation 2 predicted indifference points of 20.5 s, 17.4 s, and 10.0 s for the fixed, rectangular, and exponential conditions, respectively. The group means from these three conditions, 24.0 s, 20.4 s, and 9.7 s, respectively, were quite consistent with the pattern predicted by Equation 2. The general bias toward the adjusting key might explain why the group means were slightly longer than predicted in two of the three cases.

Figure 4 compares the results from fixed conditions with no ILIs to those of the fixed conditions with ILIs of 4 s and 8 s. Each ILI duration was tested with both white and green houselights, in separate conditions. Because there were always three ILIs before each standard reinforcer, the 4-s ILIs added a total of 12 s to the time between a choice response and reinforcement, and the 8-s ILIs added a total of 24 s. Therefore, if the subjects responded on the basis of the total time to reinforcement, the indifference points should have increased by 12 s with the 4-s ILIs and by 24 s with the 8-s ILIs. Figure 4 shows that the indifference points generally increased in conditions with ILIs, but in most cases the increases were not as large as the 12-s or 24-s increases in the standard alternative. A one-way repeated-measures analysis of variance conducted on the results shown in Figure 4 found a statistically significant effect of condition, $F(4, 12) = 7.67$, $p < .01$. Planned comparisons showed that the indifference points with no ILI were significantly shorter than with either duration of green ILI: compared to the 4-s green ILI, $F(1, 12) = 28.12$, $p < .001$, and compared to the 8-s green ILI, $F(1, 12) = 30.92$, $p < .001$. Similarly, the indifference points with no ILI were significantly shorter than with either duration of white ILI: compared to the 4-s white ILI, $F(1, 12) = 5.73$, $p < .05$, and compared to the 8-s white ILI, $F(1, 12) = 14.18$, $p < .01$.

Despite these statistically significant results, it is clear from Figure 4 that the white ILIs had different effects on different subjects. Subjects 1 and 2 showed little or no increases in indifference points when the white ILIs were added, whereas Subjects 3 and 4 showed substantial increases in these conditions. With the 8-s white ILIs, the largest increase for any individual subject (16.2 s for Bird 4) was still less than the 24-s increase in time to reinforcement that the ILIs added to the standard alternative. With the green ILIs, the indifference points increased for all subjects, and these increases were larger in every case than with the white ILIs of the same duration. With 4-s green ILIs, the mean indifference point was 11.4 s longer than with no ILI, nearly equaling the 12-s increase in time to reinforcement for the standard alternative. However, with 8-s green ILIs, the mean indifference point was only 16.0 s longer than with no ILI, whereas the ILIs added 24 s to the standard alternative in this condition.

Condition 3 added 4-s white ILIs to the exponential schedule. In this case, the ILIs produced little or no increases in the indifference points—the mean indifference point was 10.2 s, compared to a mean of 9.7 s with the exponential schedule and no ILIs. This result is consistent with Mazur's (1989) finding that with probabilistic reinforcers, the time between trials had no effect on indifference points if white houselights were present (rather than

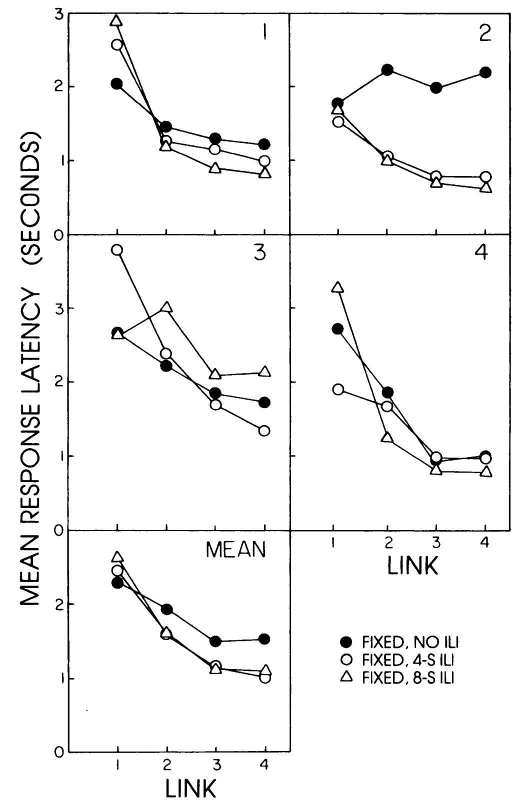

Fig. 5. Mean response latencies for each successive green-key link are shown for each subject and for the group, from the fixed conditions with no ILI or with white ILIs.

a color associated with the probabilistic reinforcer).

To determine how the different schedules on the standard key affected the subjects' rates of responding, key-peck latency—the time between the onset of the green keylight and an effective peck on that key—was recorded for each successive link before reinforcement. Figure 5 shows the mean key-peck latencies for each subject and each link in fixed conditions with no ILIs or with white ILIs. If these latencies were related to the time before reinforcement, they should have decreased across successive links, because with the fixed schedules there were always exactly four links before reinforcement; thus, each link brought the subject closer to reinforcement by a specific amount. Figure 5 shows that with only one major exception (Bird 2 in the no-ILI conditions), green-key response latencies did indeed decrease across the four links. Response latencies are not shown for the fixed conditions with

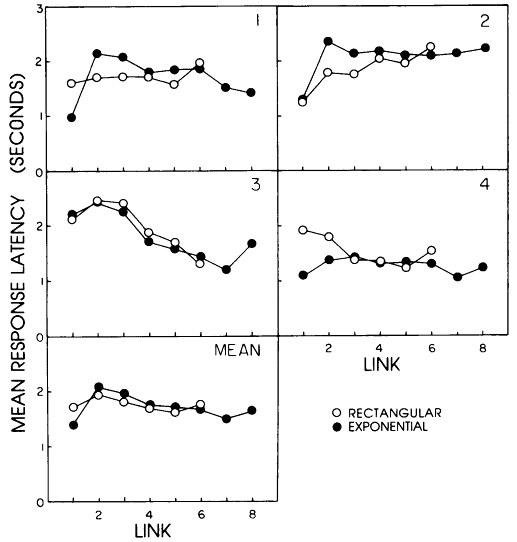

Fig. 6. Mean response latencies for each successive green-key link are shown for each subject and for the group, from the conditions with exponential and rectangular distributions of links.

green ILIs, but these latencies also decreased across links in patterns similar to those in Figure 5.

Figure 6 shows the mean green-key response latencies for the conditions with exponential and rectangular distributions of links. In these conditions, the exact number of links before reinforcement was unpredictable, ranging from one to eight with the exponential distribution and from two to six with the rectangular distribution. Subjects' response latencies reflected this unpredictability. The patterns in Figure 6 differ among subjects, but in general the functions are much flatter than in Figure 5, with little or no decrease in latencies across successive links. In several cases, most noticeably for Subject 2, latencies were shortest for the first green-key link. These shorter latencies might have occurred because in the first link, the green key was illuminated immediately after a peck on the center key, so the subject's head was not far from the green key. No peck on the center key was required to start the second or subsequent links, so the subject could be in any part of the chamber (and be facing in any direction) when the green key was illuminated. It should be noted, however, that a center key peck also preceded the first link in the fixed conditions, and the long first-link latencies in Figure 5 indicate that a required peck on the center key did not, in all conditions, lead to a prompt response on the green key. A comparison of Figures 5 and 6 suggests that latencies were relatively short if there was some probability that the current link might end in reinforcement and were relatively long if several links without reinforcement were certain to occur before reinforcement.

DISCUSSION

The results of this experiment are consistent with the theory of Rachlin et al. (1986) that probabilistic reinforcers can be viewed as delayed reinforcers, and more specifically with the position of Mazur (1991) that probabilistic reinforcers are analogous to reinforcers delivered after variable delays. The conditions with an exponential distribution of links before reinforcement were the most similar to a true probabilistic schedule, in which each response has an equal probability of delivering a reinforcer. Consistent with previous experiments with probabilistic reinforcers (Mazur, 1989, 1991), the indifference points in these conditions were substantially shorter than predicted by Equation 1, which, in effect, treats probabilistic reinforcers as if they were delivered after fixed delays. Assuming an average response latency of 1.5 s, Equation 1 predicts an indifference point of 20.5 s, compared to the mean obtained indifference point of 9.7 s. Equation 2, which takes the variability of this schedule into account, predicts an indifference point of 10.0 s.

The conditions with exactly four links before reinforcement were much more like a single fixed delay; as predicted by Equation 2, indifference points were much longer in these conditions. This difference between the fixed and exponential conditions is completely consistent with the prediction of Equation 2 that not only the *proportion* of reinforced links, but their *distribution*, must be taken into account. Equation 1 takes the proportion into account in the value of p, but it does not take into account the distribution of reinforced links. The results from the condition with a rectangular distribution provide more evidence in support of Equation 2. Although the number of links before reinforcement was variable in this condition, just as the number was variable in the exponential conditions, the indifference points (averaging 20.4 s) were much longer than in the exponential condition. These results show that it was not variability per se that led to the very short indifference points in the exponential conditions—the nature of the variability must be taken into account. Because Equation 2 does take the distribution of reinforced links into account, it correctly predicted that indifference points in the rectangular condition should be only slightly shorter than in the fixed conditions. Overall, then, the results from the conditions with different types of distributions were consistent with the hypothesis, as expressed in Equation 2, that the distribution of reinforcers over trials determines the value of a reinforcer that is delivered on a percentage basis.

Based on a study involving rats in a closed economy, Hastjarjo and Silberberg (1992) challenged the view that probabilistic reinforcers are analogous to delayed reinforcers. They reported that preference for a delayed reinforcer increased when the total amount of food per session was reduced, whereas in an earlier study (Hastjarjo, Silberberg, & Hursh, 1990), a reduction in the amount of food led to a decreased preference for a probabilistic reinforcer. Hastjarjo and Silberberg suggested that, taken together, these two studies showed that delayed and probabilistic reinforcers can have different effects on choice. Comparing results from two separate studies is risky, however, and there were many procedural differences between the two studies. In addition, the study by Hastjarjo et al. (1990) actually included two experiments, and preference for the probabilistic reinforcer increased in one and decreased in the other as food was reduced. It therefore seems difficult to draw any firm conclusions from these studies.

Another issue addressed by this experiment concerned how the addition of ILIs, and the color of the houselights during the ILIs, would affect the indifference points. Mazur (1989, 1991) concluded that any stimuli associated with a probabilistic alternative (i.e., the green keylight and green houselights in the present experiment) become conditioned reinforcers, and that only such stimuli contribute to D_i in Equation 2. The idea that the strength of a conditioned reinforcer is inversely related to the delay before primary reinforcement has many precedents (e.g., Fantino, 1969, 1977; Shull & Spear, 1987; Vaughan, 1985). It follows from Mazur's analysis that the green and white houselights presented during the ILIs should have completely different effects: With the green houselights, ILI time should be included in D_i, but with white houselights, it should not. In other words, the presence of green ILIs should have caused equally long increases in the adjusting delays, whereas the presence of white houselights should have produced no changes in the indifference points. These predictions were only partially supported. The green ILIs did produce longer indifference points. With three 4-s green ILIs, the mean indifference point increased by 11 s, close to the predicted increase of 12 s. However, with three 8-s ILIs, the mean indifference point increased by 16 s, not the 24 s predicted by Equation 2. In both cases these increases were statistically significant, but there was considerable variability across subjects. Overall, however, the increases in the indifference points were not quite as large as predicted by Equation 2. One possible explanation of this finding is that subjects may have developed a conditional discrimination in which the later portions of the green houselights were discriminated from the earlier portions. Although a 4-s green houselight was sometimes followed by food, no green houselight lasting longer than 4 s was ever followed by food. The formation of such a conditional discrimination could explain why the longer green ILIs did not have as much effect on the indifference points as predicted by Equation 2 (cf. Dunn, Williams, & Royalty, 1987; Mazur, 1991).

Contrary to the predictions of Equation 2, the inclusion of white ILIs did have an effect on the indifference points, at least for Subjects 3 and 4 (and the overall effects of the white ILIs were also statistically significant). These increases in the indifference points with white ILIs (for some subjects) are interesting not only because they violate the above interpretation of Equation 2, but also because they are different from previous results. Mazur (1989, 1991) found no evidence that the time between trials had any effect on pigeons' choices if white houselights were present. At least two aspects of the present procedure may have contributed to the different results. First, after one choice of the green key, subjects received repeated forced choices of the green key (the successive links of the standard schedule) until a reinforcer was delivered. It may be that this procedure increased the likelihood that subjects would respond to the white houselights as if they were additional links in the long chain between the initial choice response and reinforcement. Other studies with chained schedules have found that preference decreases as the number or duration of links increases (Duncan & Fantino, 1972; Leung & Winton, 1986, 1988). Another difference from the previous studies is that in those conditions in which white ILIs made a difference, the reinforcer was always delivered after the fourth link instead of on a variable basis. (In Condition 3, in which 4-s white ILIs were added to the exponential schedule, they had no effect on the indifference points.)

Without additional research, we can offer only speculation, not explanation, of why the white ILIs produced longer indifference points for some subjects. It is possible that these subjects learned to discriminate between the 4-s white houselights that occurred between green links and the much longer periods of white houselights that occurred between trials. The latter would presumably not serve as conditioned reinforcers because they were associated with the absence of reinforcement. The 4-s white houselights might have served as weak conditioned reinforcers because they were always eventually followed by the probabilistic reinforcer. Of course, the presence of individual differences poses a problem for attempts to use Equation 2 to predict the effects of the time between trials (or between links), but any other model of probabilistic reinforcement would have to confront this same problem.

Regardless of why these individual differences occurred, it is noteworthy that for all subjects, the white ILIs had smaller effects than the green ILIs. In other words, even though the total time to reinforcement was the same in both cases, all subjects showed a greater preference for the standard alternative when the houselights were white than when they were green. This finding suggests that one way to increase a pigeon's preference for a long delay before reinforcement is to present a stimulus not normally associated with the reinforcer during parts of the delay. Further research might determine exactly how such stimuli can be presented so as to maximize preference for a long reinforcer delay. Additional research might also determine whether similar effects can be observed with humans, whose responses seem to be more sensitive to the time between trials (e.g., Rachlin et al., 1986) than do those of pigeons.

REFERENCES

Cicerone, R. A. (1976). Preference for mixed versus constant delay of reinforcement. *Journal of the Experimental Analysis of Behavior,* **25,** 257–261.

Duncan, B., & Fantino, E. (1972). The psychological distance to reward. *Journal of the Experimental Analysis of Behavior,* **18,** 23–34.

Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior,* **53,** 201–218.

Dunn, R., Williams, B., & Royalty, P. (1987). Devaluation of stimuli contingent on choice: Evidence for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior,* **48,** 117–131.

Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior,* **12,** 723–730.

Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313–339). Englewood Cliffs, NJ: Prentice-Hall.

Hastjarjo, T., & Silberberg, A. (1992). Effects of reinforcer delays on choice as a function of income level. *Journal of the Experimental Analysis of Behavior,* **57,** 119–125.

Hastjarjo, T., Silberberg, A., & Hursh, S. R. (1990). Risky choice as a function of amount and variance in food supply. *Journal of the Experimental Analysis of Behavior,* **53,** 155–161.

Leung, J.-P., & Winton, A. S. W. (1986). Preference for less segmented fixed-time components in concurrent-chain schedules of reinforcement. *Journal of the Experimental Analysis of Behavior,* **46,** 175–183.

Leung, J.-P., & Winton, A. S. W. (1988). Preference for simple interval schedules of reinforcement in concurrent chains: Effects of segmentation ratio. *Journal of the Experimental Analysis of Behavior,* **49,** 9–20.

Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *Journal of Experimental Psychology: Animal Behavior Processes,* **10,** 426–436.

Mazur, J. E. (1986). Fixed and variable ratios and delays: Further tests of an equivalence rule. *Journal of Experimental Psychology: Animal Behavior Processes,* **12,** 116–124.

Mazur, J. E. (1989). Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior,* **51,** 87–99.

Mazur, J. E. (1991). Choice with probabilistic reinforcement: Effects of delay and conditioned reinforcers. *Journal of the Experimental Analysis of Behavior,* **55,** 63–77.

Mazur, J. E., Snyderman, M., & Coe, D. (1985). Influences of delay and rate of reinforcement on discrete-trial choice. *Journal of Experimental Psychology: Animal Behavior Processes,* **11,** 565–575.

Rachlin, H., Logue, A. W., Gibbon, J., & Frankel, M. (1986). Cognition and behavior in studies of choice. *Psychological Review,* **93,** 33–45.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior,* **55,** 233–244.

Rider, D. P. (1983). Preference for mixed versus constant delays of reinforcement: Effect of probability of the short, mixed delay. *Journal of the Experimental Analysis of Behavior,* **39,** 257–266.

Shull, R. L., & Spear, D. J. (1987). Detention time after reinforcement: Effects due to delay of reinforcement? In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 187–204). Hillsdale, NJ: Erlbaum.

Spetch, M. L., Belke, T. W., Barnet, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure: Effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior,* **53,** 219–234.

Vaughan, W., Jr. (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior,* **43,** 383–405.

Received January 21, 1992
Final acceptance June 7, 1992